

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2005, 16:08:55 ; Search time 72.962 Seconds  
(without alignments)  
2014.322 Million cell updates/sec

Title: US-10-003-356-5  
Perfect score: 1986  
Sequence: 1 LPHSVCTDVCPPTGTRGFVQ.....TVSTVLDLDRVLIYMCPLKIQ 380

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1986	100.0	380	5	AAE24049	Aae24049 Human V2
2	1986	100.0	927	5	AAE24050	Aae24050 Chimeric
3	1906	96.0	755	7	ADC85997	Adc85997 Human GPC
4	1749	88.1	365	5	ABP95621	Abp95621 Human GPC
5	1727.5	87.0	720	7	ADC12754	Adc12754 Human GPC
6	1444	72.7	912	8	ADI41024	Adi41024 Mouse phe
7	796.5	40.1	835	4	AA72614	Aay72614 Carassius
8	787.5	39.7	848	8	ADI40974	Adi40974 Goldfish
9	787.5	39.7	848	8	ADI41018	Adi41018 Goldfish
10	783.5	39.5	851	4	AA72617	Aay72617 Carassius
11	771.5	38.8	856	4	AA72615	Aay72615 Carassius
12	769.5	38.7	844	8	ADI41017	Adi41017 Goldfish
13	769.5	38.7	844	8	ADI40973	Adi40973 Goldfish
14	760.5	38.3	854	4	AA72616	Aay72616 Carassius
15	753.5	37.9	880	8	ADI40971	Adi40971 Fugu pher
16	753.5	37.9	880	8	ADI41021	Adi41021 Fugu pher
17	751	37.8	875	8	ADI40972	Adi40972 Fugu pher
18	751	37.8	875	8	ADI41023	Adi41023 Fugu pher
19	747	37.6	856	8	ADI41019	Adi41019 Fugu pher
20	747	37.6	856	8	ADI40970	Adi40970 Fugu pher
21	736.5	37.1	1026	2	AAW32059	Aaw32059 Dogfish s
22	736.5	37.1	1027	5	AAU76004	Aau76004 Shark kid
23	736.5	37.1	1027	5	ABB78761	Abb78761 Dogfish s
24	736.5	37.1	1027	7	ADH10917	Adh10917 Shark pol
25	736.5	37.1	1027	7	ABW02706	Abw02706 Dogfish s

26	736.5	37.1	1027	8	ADI19970	Adi19970 Dogfish s
27	722.5	36.4	1059	4	AAU00508	Aau00508 Chicken c
28	722	36.4	864	8	ADI40969	Adi40969 Fugu pher
29	722	36.4	864	8	ADI41022	Adi41022 Fugu pher
30	714.5	36.0	901	3	AA45001	Aay45001 Human cal
31	714.5	36.0	1079	2	AAW54847	Aaw54847 Rat kidne
32	714.5	36.0	1079	2	AAW38275	Aaw38275 Rat kidne
33	714.5	36.0	1079	2	AAW94928	Aaw94928 Rat kidne
34	714.5	36.0	1079	2	AA41781	Aay41781 Rat parat
35	714.5	36.0	1079	2	AAW89566	Aaw89566 Rat parat
36	714.5	36.0	1079	3	AAV51828	Aay51828 Rat calci
37	714.5	36.0	1079	5	AA47823	Aab47823 RakCar3A.
38	714.5	36.0	1079	7	ADB62141	Ade62141 Rat Prote
39	714.5	36.0	1079	7	ADE62145	Ade62145 Rat Prote
40	714.5	36.0	1079	8	ADI41013	Adi41013 Rat GPCR
41	714.5	36.0	1079	8	ADI40964	Adi40964 Rat GPCR
42	714.5	36.0	1079	8	ADI41012	Adi41012 Mouse GPC
43	714.5	36.0	1079	8	ADI40963	Adi40963 Mouse GPC
44	714.5	36.0	1079	8	ADM47115	Adm47115 Rat calci
45	714.5	36.0	1079	8	ADO29212	Ado29212 Mouse GPC

ALIGNMENTS

RESULT 1  
AAE24049  
ID AAE24049 standard; protein; 380 AA.  
XX

AC AAE24049;

DT 04-OCT-2002 (first entry)

DE Human V2 vomeronasal receptor (Zvn2R1) C-terminal protein.

KW Human; V2 vomeronasal receptor; Zvn2R1; educational tool; gene therapy; receptor.  
XX

OS Homo sapiens.

Key	Location/Qualifiers
Domain	75..100
Domain	/note= "Transmembrane domain-1"
Domain	101..113
Domain	/note= "Intracellular domain"
Domain	114..134
Domain	/note= "Transmembrane domain-2"
Domain	135..145
Domain	/note= "Extracellular domain"
Domain	146..170
Domain	/note= "Transmembrane domain-3"
Domain	171..188
Domain	/note= "Intracellular domain"
Domain	189..208
Domain	/note= "Transmembrane domain-4"
Domain	209..230
Domain	/note= "Extracellular domain"
Domain	231..255
Domain	/note= "Transmembrane domain-5"
Domain	256..268
Domain	/note= "Intracellular domain"
Domain	269..289
Domain	/note= "Transmembrane domain-6"
Domain	290..330
Domain	/note= "Extracellular domain"
Domain	326..380
Domain	/note= "Intracellular domain"
Domain	331..325
Domain	/note= "Transmembrane domain-7"

XX WO200242464-A2.

PN 30-MAY-2002.

PD

XX PF 15-NOV-2001; 2001WO-US046034.  
XX PR 21-NOV-2000; 2000US-0252373P.  
XX PA (ZYMO ) ZYMOGENETICS INC.  
XX PI Lok S, Holloway JL;  
XX DR WPI; 2002-479953/51.  
XX DR N-PSDB; AAD39170.  
XX PT Novel isolated human V2 vomeronasal receptor, termed Zvn2R1, for  
PT identifying presence of Zvn2R1 ligand in sample, as educational tools in  
PT laboratory practicum kits for courses related to genetics and molecular  
PT biology.  
XX PS Claim 1; Page 85-86; 98pp; English.  
XX CC The invention relates to an isolated human V2 vomeronasal receptor termed  
CC Zvn2R1. The Zvn2R1 nucleic acid is useful for detecting the expression of  
CC Zvn2R1 gene in a biological sample, to determine if a subject's  
CC chromosomes contain a mutation in the Zvn2R1 gene, and for therapeutic  
CC purposes. Zvn2R1 is useful as an aid to teach preparation of antibodies,  
CC identifying proteins by Western blotting, protein purification,  
CC determining the weight of expressed Zvn2R1 polypeptides as a ratio to  
CC total protein expressed, identifying peptide cleavage sites, coupling  
CC amino and carboxyl terminal tags, amino acid sequence analysis,  
CC monitoring biological activities of both the native and tagged protein in  
CC vitro and in vivo and to teach analytical skills such as mass  
CC spectrometry, circular dichroism to determine conformation, especially of  
CC the four alpha helices X-ray crystallography to determine the three-  
CC dimensional structure in atomic detail, and nuclear magnetic resonance  
CC spectroscopy to reveal the structure of proteins in solution. Zvn2R1 is  
CC useful as educational tools in laboratory practicum kits for courses  
CC related to genetics and molecular biology, protein chemistry, antibody  
CC production and analysis, and as standards or as unknowns for testing  
CC purposes. The invention is useful as a teaching aid to instruct students  
CC how to prepare affinity chromatography columns to purify Zvn2R1, and for  
CC cloning and sequencing the polynucleotide that encodes an antibody and  
CC thus as a practicum for teaching a student how to design humanised  
CC antibodies. The invention is useful in gene therapy. The present sequence  
CC is human Zvn2R1 C-terminal protein  
XX SQ Sequence 380 AA;  
Query Match 100.0%; Score 1986; DB 5; Length 380;  
Best Local Similarity 100.0%; Pred. No. 4.6e-201;  
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LPHSVCTDVCPPGTRGTFVQREPICCFDSIPCADGHVSRKPGBRECEQCGEDYWSNAQS 60  
Db 1 LPHSVCTDVCPPGTRGTFVQREPICCFDSIPCADGHVSRKPGBRECEQCGEDYWSNAQS 60  
QY 61 ECVLKEVEYLAYDEALGFTLVILSVFGAFVVLAVTAVYVIHRHTPLVNASDWQLGFLIQV 120  
Db 61 ECVLKEVEYLAYDEALGFTLVILSVFGAFVVLAVTAVYVIHRHTPLVNASDWQLGFLIQV 120  
QY 121 SLIIMLLSSMLFIDKPHNWSMAGQVTLALGFSCLSLCKTSSFLAYRISKSTQLT 180  
Db 121 SLIIMLLSSMLFIDKPHNWSMAGQVTLALGFSCLSLCKTSSFLAYRISKSTQLT 180  
QY 181 SMHPLYRKIIIVLSVLAIEIGICTAYLILEPPMVYKNMESQNTKIILGCNEISIEFLYSMF 240  
Db 181 SMHPLYRKIIIVLSVLAIEIGICTAYLILEPPMVYKNMESQNTKIILGCNEISIEFLYSMF 240  
QY 241 GIDAFALLCFLTTFVARQLPDNYYEGKCITFGMLVFFIIVMSFVPVYLSTKGKFMAVE 300  
Db 241 GIDAFALLCFLTTFVARQLPDNYYEGKCITFGMLVFFIIVMSFVPVYLSTKGKFMAVE 300  
QY 301 IFAILASSHGLGICIFAPKCLIIILLRPERNTSEIVCGRVSTTDNCIQLTSAFVSSELNNT 360  
Db 301 IFAILASSHGLGICIFAPKCLIIILLRPERNTSEIVCGRVSTTDNCIQLTSAFVSSELNNT 360

QY 361 TVSTVLDDRVLIYMCPLKLQ 380  
Db 361 TVSTVLDDRVLIYMCPLKLQ 380  
RESULT 2  
AAE24050  
ID AAE24050 standard; protein; 927 AA.  
XX AC AAE24050;  
XX 29-AUG-2003 (revised)  
DT 04-OCT-2002 (first entry)  
XX Chimeric receptor DNA protein.  
DE Human; V2 vomeronasal receptor; Zvn2R1; educational tool; gene therapy;  
XX receptor; murine; chimeric.  
KW Homo sapiens.  
XX Mus sp.  
OS Chimeric.  
FH Key Location/Qualifiers  
FT Domain 1. .621 /note= "Extracellular domain"  
FT Peptide 1. .29 /label= Signal\_peptide  
FT Protein 30. .927 /note= "Mature chimeric receptor protein"  
FT Domain 30. .610 /note= "Ligand binding domain"  
FT Domain 622. .647 /note= "Transmembrane domain-1"  
FT Domain 648. .660 /note= "Intracellular domain"  
FT Domain 661. .681 /note= "Transmembrane domain-2"  
FT Domain 682. .692 /note= "Extracellular domain"  
FT Domain 693. .717 /note= "Transmembrane domain-3"  
FT Domain 718. .735 /note= "Intracellular domain"  
FT Domain 736. .755 /note= "Transmembrane domain-4"  
FT Domain 756. .777 /note= "Extracellular domain"  
FT Domain 778. .802 /note= "Transmembrane domain-5"  
FT Domain 803. .815 /note= "Intracellular domain"  
FT Domain 816. .836 /note= "Transmembrane domain-6"  
FT Domain 837. .847 /note= "Extracellular domain"  
FT Domain 848. .872 /note= "Transmembrane domain-7"  
FT Domain 873. .927 /note= "Intracellular domain"  
XX WO200242464-A2.  
PN 30-MAY-2002.  
XX 15-NOV-2001; 2001WO-US046034.  
XX 21-NOV-2000; 2000US-0252373P.  
XX (ZYMO ) ZYMOGENETICS INC.  
PI Lok S, Holloway JL;

XX WPI; 2002-479953/51.  
DR N-PSDB; AAD39172.  
XX  
PT Novel isolated human V2 vomeronasal receptor, termed Zvn2R1, for  
PT identifying presence of Zvn2R1 ligand in sample, as educational tools in  
PT laboratory practicum kits for courses related to genetics and molecular  
PT biology.  
XX  
PS Claim 5; Page 93-96; 98pp; English.  
XX  
CC The invention relates to an isolated human V2 vomeronasal receptor termed  
CC Zvn2R1. The Zvn2R1 nucleic acid is useful for detecting the expression of  
CC Zvn2R1 gene in a biological sample, to determine if a subject's  
CC chromosomes contain a mutation in the Zvn2R1 gene, and for therapeutic  
CC purposes. Zvn2R1 is useful as an aid to teach preparation of antibodies,  
CC identifying proteins by Western blotting, protein purification,  
CC determining the weight of expressed Zvn2R1 polypeptides as a ratio to  
CC total protein expressed, identifying peptide cleavage sites, coupling  
CC amino and carboxyl terminal tags, amino acid sequence analysis,  
CC monitoring biological activities of both the native and tagged protein in  
CC vitro and in vivo and to teach analytical skills such as mass  
CC spectrometry, circular dichroism to determine conformation, especially of  
CC the four alpha helices X-ray crystallography to determine the three-  
CC dimensional structure in atomic detail, and nuclear magnetic resonance  
CC spectroscopy to reveal the structure of proteins in solution. Zvn2R1 is  
CC useful as educational tools in laboratory practicum kits for courses  
CC related to genetics and molecular biology, protein chemistry, antibody  
CC production and analysis, and as standards or as unknowns for testing  
CC purposes. The invention is useful as a teaching aid to instruct students  
CC how to prepare affinity chromatography columns to purify Zvn2R1, and for  
CC cloning and sequencing the polynucleotide that encodes an antibody and  
CC thus as a practicum for teaching a student how to design humanised  
CC antibodies. The invention is useful in gene therapy. The present sequence  
CC is chimeric receptor protein. This chimeric sequence was designed by  
CC aligning human Zvn2R1 and murine tissue-type vomeronasal putative  
CC pheromone receptor (V2R2). (Updated on 29-AUG-2003 to standardise OS  
CC field)  
XX  
SQ Sequence 927 AA;

Query Match 100.0%; Score 1986; DB 5; Length 927;  
Best Local Similarity 100.0%; Pred. No. 1.5e-200;  
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LPHSVCTDVCPPGTGRGVQREPICCFDSIPCADGHVSRKPGERECEQCCEGYWSNAQKS 60  
Db 548 LPHSVCTDVCPPGTGRGVQREPICCFDSIPCADGHVSRKPGERECEQCCEGYWSNAQKS 607  
QY 61 ECVLKEVEYLAYDEALGFTLVILSVFGAFVVLAVTAVYVIHRHTPLVNASDWQLGFLIQV 120  
Db 608 ECVLKEVEYLAYDEALGFTLVILSVFGAFVVLAVTAVYVIHRHTPLVNASDWQLGFLIQV 667  
QY 121 SLIIMLLSSMLFIDKPHNWSMAGQVTLALGFSCLCLLGKTSLSFLAYRISKSTQLT 180  
Db 668 SLIIMLLSSMLFIDKPHNWSMAGQVTLALGFSCLCLLGKTSLSFLAYRISKSTQLT 727  
QY 181 SMHPLYRKIIIVLSVLAIEIGICTAYLILEPPMVYKNMESQNTKIIILGCNEISIEFLYSMF 240  
Db 728 SMHPLYRKIIIVLSVLAIEIGICTAYLILEPPMVYKNMESQNTKIIILGCNEISIEFLYSMF 787  
QY 241 GIDAFALLCFLTTFVARQLPDNYYEGKCITFGMLVFFIWMSPVYVYSTKGKFMAVE 300  
Db 788 GIDAFALLCFLTTFVARQLPDNYYEGKCITFGMLVFFIWMSPVYVYSTKGKFMAVE 847  
QY 301 IFAILASSHGLGCIAPKCLIIILLRPERNTSEIVCGRVSTTDNCIQLTSAFVSSELNNT 360  
Db 848 IFAILASSHGLGCIAPKCLIIILLRPERNTSEIVCGRVSTTDNCIQLTSAFVSSELNNT 907  
QY 361 TVSTVLDLDRVLIYMCPLKIQ 380  
Db 908 TVSTVLDLDRVLIYMCPLKIQ 927

RESULT 3  
ADC85997  
ID ADC85997 standard; protein; 755 AA.  
XX  
AC ADC85997;  
XX  
DT 01-JAN-2004 (first entry)  
XX  
DE Human GPCR protein SEQ ID NO:450.  
XX  
KW human; GPCR; guanosine triphosphate-binding protein coupled receptor;  
KW gene therapy.  
XX  
OS Homo sapiens.  
XX  
PN EP1270724-A2.  
XX  
PD 02-JAN-2003.  
XX  
PF 18-JUN-2002; 2002EP-00013517.  
XX  
PR 18-JUN-2001; 2001JP-00246789.  
XX  
PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.  
PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.  
XX  
PI Suwa M, Asai K, Akiyama Y, Aburatani H;  
XX  
DR WPI; 2003-315783/31.  
DR N-PSDB; ADC85996.  
XX  
PT New polynucleotide, useful for preparing a composition for treating a  
PT patient in need of increased or suppressed activity or expression of the  
PT guanosine triphosphate-binding protein coupled receptor.  
XX  
PS Claim 2; SEQ ID NO 450; 28pp; English.  
XX  
CC The invention relates to a novel polynucleotide encoding a guanosine  
CC triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of  
CC the invention may have a use in gene therapy. The polynucleotide and  
CC polypeptide are useful for preparing a composition for treating a patient  
CC in need of increased or suppressed activity or expression of the  
CC guanosine triphosphate-binding protein coupled receptor. The protein  
CC sequences shown in ADC85549-ADC87617 represent GPCR's of the invention.  
XX  
SQ Sequence 755 AA;  
Query Match 96.0%; Score 1906; DB 7; Length 755;  
Best Local Similarity 99.7%; Pred. No. 3.4e-192;  
Matches 367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 13 GTGRGFVQREPICCFDSIPCADGHVSRKPGERECEQCCEGYWSNAQKSECVLKEVEYLAY 72  
Db 388 GLGRGFVQREPICCFDSIPCADGHVSRKPGERECEQCCEGYWSNAQKSECVLKEVEYLAY 447  
QY 73 DEALGFTLVILSVFGAFVVLAVTAVYVIHRHTPLVNASDWQLGFLIQVSLIIMLLSSMLF 132  
Db 448 DEALGFTLVILSVFGAFVVLAVTAVYVIHRHTPLVNASDWQLGFLIQVSLIIMLLSSMLF 507  
QY 133 IDKPHNWSMAGQVTLALGFSCLCLLGKTSLSFLAYRISKSTQLTSMHPLYRKIIIVL 192  
Db 508 IDKPHNWSMAGQVTLALGFSCLCLLGKTSLSFLAYRISKSTQLTSMHPLYRKIIIVL 567  
QY 193 ISVLAIEIGICTAYLILEPPMVYKNMESQNTKIIILGCNEISIEFLYSMFIDAFALLCFL 252  
Db 568 ISVLAIEIGICTAYLILEPPMVYKNMESQNTKIIILGCNEISIEFLYSMFIDAFALLCFL 627  
QY 253 TTFVARQLPDNYYEGKCITFGMLVFFIWMSPVYVYSTKGKFMAVEIFAILASSHGLL 312  
Db 628 TTFVARQLPDNYYEGKCITFGMLVFFIWMSPVYVYSTKGKFMAVEIFAILASSHGLL 687  
QY 313 GCIFAPKCLIIILLRPERNTSEIVCGRVSTTDNCIQLTSAFVSSELNNTTVSTVLDLDRVLI 372



Db 688 GCIFAPKCLIIILRPERNTSEIVCGRVSTTDNCIQLTSAFVSSELNNTTSTVLDLDRVLI 747  
QY 373 YMCPLKLQ 380  
Db 748 YMCPLKLQ 755  
RESULT 4  
ABP95621 ID ABP95621 standard; protein; 365 AA.  
AC ABP95621;  
XX  
DT 06-MAR-2003 (first entry)  
XX  
DE Human GPCR polypeptide SEQ ID NO 52.  
XX  
KW Human; GPCR; G protein coupled receptor; signal transduction; olfactory;  
KW drug development; gustatory; taste; fragrance; receptor.  
XX  
OS Homo sapiens.  
XX  
PN WO200216548-A2.  
XX  
PD 28-FEB-2002.  
XX  
PF 30-JUL-2001; 2001WO-IB001446.  
XX  
PR 04-AUG-2000; 2000JP-00237818.  
PR 13-FEB-2001; 2001JP-00034434.  
XX  
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.  
XX  
PI Haga T, Takeda S, Mitaku S;  
XX  
DR WPI; 2002-304118/34.  
DR N-PSDB; ABZ42895.  
XX  
PT Database global search for G protein-coupled receptors, proteins and  
PT encoded genes for studying in vivo signal transduction mechanism and  
PT identifying targets for drug development.  
XX  
PS Claim 10; SEQ ID NO 52; 97pp + Sequence Listing; Japanese.  
XX  
CC The invention relates to a method for screening G protein-coupled  
CC receptor (GPCR) genes (ABZ42870-ABZ43216) and/or GPCR proteins (ABP95596-  
CC ABP95942) by extracting open-reading frames containing 6-8 transmembrane  
CC domains with 250-1000 amino acid residues to give a gene homologous with  
CC a known GPCR gene. The receptor proteins and encoded genes are useful for  
CC studying in vivo signal transduction mechanism and identifying targets  
CC for drug development e.g. based on olfactory and gustatory receptors in  
CC form of agonists and antagonists by screening intrinsic and extrinsic  
CC ligands as bitter taste inhibitors, taste enhancers and fragrance  
CC improvers. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
SQ Sequence 365 AA;  
Query Match 88.1%; Score 1749; DB 5; Length 365;  
Best Local Similarity 99.1%; Pred. No. 5.4e-176;  
Matches 341; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 37 VSRKPGERECEQCGEDYWSNAQSECVLKEVEYLAYDEALGFTLVILSVFGAFVVLAVTA 96  
Db 22 VLRSIGERECEQCGEDYWSNAQSECVLKEVEYLAYDEALGFTLVILSVFGAFVVLAVTA 81  
QY 97 VYVIHRHTPLVNASDWQLGFLIQVSLIIMLLSSMLFTDKPHNWSMAGQVTLALGFSLCL 156  
Db 82 VYVIHRHTPLVNASDWQLGFLIQVSLIIMLLSSMLFTDKPHNWSMAGQVTLALGFSLCL 141  
QY 157 SCLLGKTSLSFLAYRISKSTQLTSMHPLYRKIIIVLSVLAIEIGICTAYLILEPPMVYKN 216

Db 142 SCLLGKTSLSFLAYRISKSTQLTSMHPLYRKIIIVLSVLAIEIGICTAYLILEPPMVYKN 201  
QY 217 MESQNTKIILGCNEISIEFLYSFMFGIDAFLLALLCFLTTTFVARQLPDNYEGKCTTFGMLV 276  
Db 202 MESQNTKIILGCNEISIEFLYSFMFGIDAFLLALLCFLTTTFVARQLPDNYEGKCTTFGMLV 261  
QY 277 FFIIMWSFVVPVYLSTKGKFMAVEIFAILASSHGLLGCFAPKCLIIILRPERNTSEIVC 336  
Db 262 FFIIMWSFVVPVYLSTKGKFMAVEIFAILASSHGLLGCFAPKCLIIILRPERNTSEIVC 321  
QY 337 GRVSTTDNCIQLTSAFVSSELNNTTSTVLDLDRVLIYMCPLKLQ 380  
Db 322 GRVSTTDNCIQLTSAFVSSELNNTTSTVLDLDRVLIYMCPLKLQ 365  
RESULT 5  
ADC12754 ID ADC12754 standard; protein; 720 AA.  
XX  
AC ADC12754;  
XX  
DT 18-DEC-2003 (first entry)  
XX  
DE Human GPCR protein, SEQ ID NO 86.  
XX  
KW G protein-coupled receptor; GPCR; antibacterial; fungicide; protozoacide;  
KW virucide; antirheumatic; antiarthritic; tranquiliser; antidiabetic;  
KW osteopathic; nootropic; neuroprotective; anorectic; cardiant;  
KW neuroleptic; cytostatic; antiparkinsonian; hypotensive; hypertensive;  
KW antiulcer; antiallergic; anticonvulsant; analgesic; infection;  
KW rheumatoid arthritis; chronic obstructive pulmonary diseases; COPD;  
KW asthma; non-insulin dependent diabetes; obesity; osteoporosis;  
KW Alzheimer's disease; age-related macular degeneration;  
KW myocardial infarction; schizophrenia; osteoarthritis; cancer;  
KW Parkinson's disease; congestive heart failure; hypotension; hypertension;  
KW ulcer; allergy; benign prostatic hyperplasia; seizure disorder; anxiety;  
KW obsessive compulsive disorder; Cushing's syndrome; hypopituitarism; pain;  
KW human.  
XX  
OS Homo sapiens.  
XX  
PN WO2003000893-A2.  
XX  
PD 03-JAN-2003.  
XX  
PF 24-JUN-2002; 2002WO-IB002357.  
XX  
PR 26-JUN-2001; 2001US-0301095P.  
PR 06-NOV-2001; 2001US-0333185P.  
XX  
PA (DECO-) DECODE GENETICS EHF.  
XX  
PI Martinez RMA, Sigurdsson GT;  
XX  
DR WPI; 2003-210155/20.  
DR N-PSDB; ADC12753.  
XX  
PT New G protein-coupled receptor (GPCR) genes and polypeptides, useful for  
PT diagnosing diseases associated with a GPCR, or in gene therapy for  
PT treating e.g. obesity, osteoporosis, Alzheimer's, cancers or congestive  
PT heart failure.  
XX  
PS Claim 10; SEQ ID NO 86; 253pp; English.  
XX  
CC The invention relates to a novel isolated nucleic acid of a G protein-  
CC coupled receptor (GPCR) gene comprising any of 62 sequences of 912-2454  
CC bp, or its complements; a GPCR polypeptide comprising any of 62 sequences  
CC of 291-818 amino acids; or a nucleic acid that hybridises, under high  
CC stringency conditions, with any of the 62 GPCR sequences or any of their  
CC complements. The GPCR agents of the invention have the following  
CC activities: antibacterial, fungicide, protozoacide, virucide,  
CC antirheumatic, tranquiliser, antiarthritic, antidiabetic, osteopathic,



CC represents a species homologue of a novel GPCR of the invention.

XX Sequence 912 AA;

Query Match 72.7%; Score 1444; DB 8; Length 912;  
Best Local Similarity 74.7%; Pred. No. 3.5e-143;  
Matches 274; Conservative 37; Mismatches 56; Indels 0; Gaps 0;

QY 2 PHSVCTDVCPPGTGRGVQREPICCFDSIPCADGHVSRKPGERECEQCCEGYWSNAQKSE 61  
Db PDSFCTQVCPGTRKGIQGPICCFDCIPCADGVSEKSGQRECDPCGEDDWSNAGKSK 605  
QY 62 CVLKEVEYLAYDEALGFTLVILSVFGAFVVLAVTAVVVIHRHTPLVNASDWQLGFLIQVS 121  
Db CVPKLVEFLAYGEALGFTLVILSIFGALVVLAVTVVVIHRHTPLVKANDRELSFLIQMS 665  
QY 122 LIIMLLSSMLFIDKPHNWSMAGQVTLALGFSCLCLLGGTSSFLAYRISKSTQLTS 181  
Db LVITVLSLLFIGKPCNWSMARQITLALGFCLCLSSILGKTIISLFFAYRISVSKTRLIS 725  
QY 182 MHPLYRKIIIVLISVLAEGICTAYLILEPPMVYKNMESQNTKIILGCNEISIEFLYSMFG 241  
Db MHPFIRKLIIVCVVGEIGVCAAYLVLEPPRMFKNIEIQNKIIFECNEGSVEFLCSIFG 785  
QY 242 IDAFLALLCFLTTTFVARQLPDNYYEGKCITFGMLVFFIIMSFVPVYLSTKGFKMAVEI 301  
Db FDLRALLCFLTTTFVARQLPDNYYEGKCITFGMLVFFIIMSFVPVYLSTKGFKMAVEI 845  
QY 302 FAILASSHGLGCFAPKCLIIILLRPERNTSEIVCGRVSTTDNCIQLTSAFVSSELNNTT 361  
Db FAILASSYGLLGLFLPKCFIILLRKRNTDTEVGRVPTVDRSIQLTSAVSSELNSTT 905  
QY 362 VSTVLDD 368  
Db VSTVLDE 912

RESULT 7  
AAV72614

ID AAY72614 standard; protein; 835 AA.

AC AAY72614;

DT 02-MAY-2001 (first entry)

XX Carassius auratus full-length CaSR-like protein #1.

XX Goldfish; G protein-coupled odorant family receptor; R5.24; screening;  
KW conception; reproductive behaviour; sexual behaviour;  
KW non-sexual social behaviour; olfactory system; reproductive physiology;  
KW feeding behaviour; migratory behaviour; implantation; oestrous;  
KW menstruation; CaSR-like protein.

XX Carassius auratus.

OS Key Location/Qualifiers

XX Misc-difference 338

FT /label= Unknown

FT Misc-difference 361

FT /label= Unknown

FT Misc-difference 379

FT /label= Unknown

FT Misc-difference 380

FT /label= Unknown

FT Misc-difference 406

FT /label= Unknown

XX WO200105833-A1.

XX 25-JAN-2001.

XX 19-JUL-2000; 2000WO-US019687.

XX

PR 20-JUL-1999; 99US-0144766P.

XX (REGC ) UNIV CALIFORNIA.

PA Ngai J, Specia DJ, Lin DM, Isacoff EY, Dittman AH, Fan J;

XX WPI; 2001-159517/16.

PI Novel G protein-coupled odorant family receptors, useful for screening  
XX compounds capable of modulating reproductive/sexual and non-sexual social  
PT behaviors.

PS Claim 6; Page 40-42; 62pp; English.

XX The patent discloses methods and compositions relating to odorant  
CC receptors, including a general expression cloning methodology which is  
CC useful for identifying novel G protein-coupled receptors and a novel  
CC family of odorant receptors and related nucleic acids, ligands, agonists  
CC and antagonists. R5.24 which is an odorant receptor and the nucleic acid  
CC sequences encoding R5.24 are useful for screening related receptors,  
CC agonists and antagonists of R5.24, which are useful for modulating  
CC reproductive/sexual and non-sexual social behaviours mediated through the  
CC olfactory system, reproductive physiologies and olfactory system  
CC regulated feeding behaviours, migratory behaviours and events such as  
CC conception, implantation, oestrous, and menstruation. R5.24 nucleic acid  
CC sequence is useful as translatable transcripts, hybridisation probes, PCR  
CC primers, diagnostic nucleic acids, for detecting the presence of R5.24  
CC genes and gene transcripts, and in detecting or amplifying nucleic acids  
CC encoding additional R5.24 homologues and structural analogues. The  
CC present sequence is Carassius auratus full-length CaSR-like protein. CaSR  
CC -like receptor family shows similarity to receptor 5.24 and is  
XX specifically expressed in the goldfish olfactory epithelium

SQ Sequence 835 AA;

Query Match 40.1%; Score 796.5; DB 4; Length 835;

Best Local Similarity 43.5%; Pred. No. 1.1e-74;

Matches 147; Conservative 72; Mismatches 118; Indels 1; Gaps 1;

QY 2 PHSVCTDVCPPGTGRGVQREPICCFDSIPCADGHVSRKPGERECEQCCEGYWSNAQKSE 61

Db PNVCSSESCLPGTRKAAQKGRPVCCYDCIPCAEGEISNETSDNNCKQCPREYWSNAEKT 555

QY 62 CVLKEVEYLAYDEALGFTLVILSVFGAFVVLAVTAVVVIHRHTPLVNASDWQLGFLIQVS 121

Db CVLKAVEFLSFTVMGIVLAFVFLGAGLTALVAILFYMRDTPIVKANNSLSFLLFS 615

QY 122 LIIMLLSSMLFIDKPHNWSMAGQVTLALGFSCLCLLGGTSSFLAYRISKSTQLTS 181

Db LTLCLCSLTFIGOPNEWSCMLRHTAFGITFVLCISCVLGTIVVLMFAKATLPGSNVMK 675

QY 182 -MHPLYRKIIIVLISVLAEGICTAYLILEPPMVYKNMESQNTKIILGCNEISIEFLYSMF 240

Db WFGPAQQRLSVLALTFIQILICVLWLTISPFPYKNMKYFKEKIILECSLGSSISFWAVL 735

QY 241 GIDAFLLALLCFLTTTFVARQLPDNYYEGKCITFGMLVFFIIMSFVPVYLSTKGFKMAVE 300

Db GYIGLLAVLCLFILAFLARTLPDNFNEAKFITFSMLIFCAVWITFIPAYVSSPGKYTVAVE 795

QY 301 IFAILASSHGLGCFAPKCLIIILLRPERNTSEIVCGR 338

Db IFAILASSFGLLFCIFAPKCYIILLKPDQNTKKHMMGK 833

RESULT 8

ADI40974

ID ADI40974 standard; protein; 848 AA.

XX AC ADI40974;

XX 22-APR-2004 (first entry)

DT Goldfish putative odorant receptor 2.

XX

DE









agonists and antagonists of R5.24, which are useful for modulating reproductive/sexual and non-sexual social behaviours mediated through the olfactory system, reproductive physiologies and olfactory system regulated feeding behaviours, migratory behaviours and events such as conception, implantation, oestrous, and menstruation. R5.24 nucleic acid sequence is useful as translatable transcripts, hybridisation probes, PCR primers, diagnostic nucleic acids, for detecting the presence of R5.24 genes and gene transcripts, and in detecting or amplifying nucleic acids encoding additional R5.24 homologues and structural analogues. The present sequence is *Carassius auratus* full-length CaSR-like protein. CaSR-like receptor family shows similarity to receptor 5.24 and is specifically expressed in the goldfish olfactory epithelium

Sequence 851 AA;

Query Match	39.5%;	Score 783.5;	DB 4;	Length 851;
Best Local Similarity	42.6%;	Pred. No. 2.6e-73;		
Matches 144;	Conservative 69;	Mismatches 124;	Indels 1;	Gaps 1;
QY	2	PHSVCTDVCPPTGRGFVQREPICPDFSIPCADGHVSRKPGERECEQC GEDYWSNAQKSE	61	
Db	513	PVSVCSESCPPGTRKAVQKGRPVCCYDCIPCSEGEINNETDSSDCFPDCL EYWSNEGDK	572	
QY	62	CVLKEVEYLAYDEALGFTLVILSVFGAFVVLAVTAVVVIHRHTPLVNASDWQLGFLIQVS	121	
Db	573	CVLKWEFLSYTEIMGTVLCIFSF FFGMLLTAVSVFVYLHKETPIVRANNSELSFLLLPS	632	
QY	122	LIIMLLSSMLFIDKPHNWSMCAGQVTLALGFSLCLSLGLGKTSLLFLAYRIS-KSKTQLT	180	
Db	633	LSLCFLCSLTFIGRPTWSCMLRHTAFGVTFVLCISCVLGKTIIVLMAFRATLPGSNVMK	692	
QY	181	SMHPLYRKIIIVLSVLAIEGICTAYLILEPPVMYKNMESQNTKIILGCNEISIEFLYSMF	240	
Db	693	CFGPLQQRFSVVSLSLTQMIICVLWLITISPPFPFNNLSYYREKIIILECNLGSALGFVGVL	752	
QY	241	GIDAFALLCFLTTFVARQLPDNYEYEGKCITFGMLVPFIWMSFVVPVYLSTKGKFKMAVE	300	
Db	753	GYTGLLSILCFILAFIARKLPDNFNEAKFITFSMLIFCAVWITFIPAYVSSPGKFTTAVQ	812	
QY	301	IFAILASSHGLGCIAPKPKLIILLRPERNTSEIVCGR	338	
Db	813	IFAILASSFSLFCIFAPKCYIILLKPEKNTKKQIMGK	850	

RESULT 11	
AAAY72615	
ID	AAAY72615 standard; protein; 856 AA.
XX	
AC	AAAY72615;
XX	
DT	02-MAY-2001 (first entry)
XX	
DE	Carassius auratus full-length CasR-like protein #2.
XX	
KW	Goldfish; G protein-coupled odorant family receptor; R5.24; screening;
KW	conception; reproductive behaviour; sexual behaviour;
KW	non-sexual social behaviour; olfactory system; reproductive physiology;
KW	feeding behaviour; migratory behaviour; implantation; oestrous;
KW	menstruation; CasR-like protein.

XX	Carassius auratus.
FH	Key
FT	Misc-difference 841
FT	/label= Unknown
FT	Misc-difference 849
FT	/label= Unknown

PN WO200105833-A1.

PD 25-JAN-2001.

19-JUL-2000: 2000WO-US019687.

**XX PR XX PA XX PI XX DR XX PT PT PT XX PS XX CC**

Query Match 38.8%; Score 771.5; DB 4; Length 856;  
Best Local Similarity 42.2%; Pred. No. 5e-72;  
Matches 141; Conservative 66; Mismatches 126; Indels 1; Gaps 1;

Qy	1	LPHSVCTDVCPPGTGRGVQREPICCFDSIPCADGHVSRKPGERECEQCGEDYWSNAQKS	60
Db	511	LPVSVCSCTCPPGTRKAVQKGRPVCCYDCIPCCEGEISNGTSDNDCFPDCLYWSNESND	570
Qy	61	ECVLKEVBYLAYDEALGFTLVILSVFGAFVVLAVTAVVYIHRHTPLVNASDWQLGFLIQV	120
Db	571	RCVLKVIKFLSYTEIMGVMVLCIPSPFIGVLLTTIVSFLFYLHKETPIVRANNSLSFLLLF	630
Qy	121	SLIIMLLSSMLFIDKPHNWSCMAGQVTLALGFSCLCLGKTSLSFLAYRISKSKTQLT	180
Db	631	SLTLCFLCSLTFIGRPTWESCMRLRHTAFGITFVLCISCILGKTIIVLMAFKATLPGSNVM	690
Qy	181	S-MHPLYRKIIIVLSVLAEIGICTAYLILEPPMVYXNMESONTKIIILGCNEISIEFLYSM	239
Db	691	KWFGPLQQQLSVVSLTLIQMIICVLWLTISPPPPYMNLSYYREKIIILECNVGSDLAFAVAV	750
Qy	240	FGIDAFALALLCFLTTFVARQLPDNYEGKCIITFGMLVFFIWNMSFVYVYLSLTKGKFKMAV	299
Db	751	LGYTGLLSILCFVLAFLARKLPDNFNEAKFITFSMLIFCAVWLTFIPAYVSSPGKFTVAV	810
Qy	300	EIFAILASSHGLLGCIFAPKCLIIILLRPERNTSE	333
Db	811	EIFAILASSFSLLCFIFAPKCYIILLKPEKQKHS	844

## RESULT 12

ADI41017

ID ADI41017 standard; protein; 844 AA.

AC ADI41017;

DT 22-APR-2004 (first entry)

XX









		Matches 139;	Conservative 69;	Mismatches 123;	Indels 1;	Gaps 1;
Qy	1	LPHSVCTDVCPPGTRGFGVQREPICCFDSIPCADGHVSRKPGERECEQCCEGYWSNAQKS	60			
Db	538	VPLSVCSSICPPGTRKAIRPNYPICCHDCVVCTAGEISNQDAIECARCLPEFWSNADRT	597			
Qy	61	ECVLKEVEYLAYDEALGFTLVILSVFGAFVVLAVTAVYVIHRHTPLVNASDWQLGFLIQV	120			
Db	598	ACVPKQVEFLSFGDTIGIALLVWSLIGSFLTCAVALVFFYHRTSPIVRANNSDLSFLLLF	657			
Qy	121	SLIIMLLSSMLFIDKPHNWSMAGQVTLALGFSCLCLLGTSSFLAYRISKSTQLT	180			
Db	658	SLTLCFLCSLTFISPPSQWSCMLRHTAFGITFVLCISILGKTIIVLMAFRATLPGSDVM	717			
Qy	181	S-MHPLYRKIIIVLISVLAIEIGICTAYLIILEPPMVYKNMESQNTKIIILGCNEISIEFLYSM	239			
Db	718	KWFGPGKQKAIITFSTLVQVWICTVWLVAAPTPTQYMPRESAIIILLCDEGSTIAFSLV	777			
Qy	240	FGIDAFLLALLCPLTTFVARQLPDNYVEGKCITFGMLVFFIIMMSFVPVYLSKGFKMAV	299			
Db	778	LGYIGVLACMCFLLAFLARKLPDNFNEARLIAFSMLIFCAVWVAFVAYISSPGKYSTLT	837			
Qy	300	EIPAILASSHGLGCIAPKCLIIILLRPERNT	331			
Db	838	EIPAILASSYGLGCIAPKCYIIILMKSEKNT	869			

Search completed: February 16, 2005, 16:32:01  
Job time : 73.962 secs

**THIS PAGE LEFT BLANK**